

M M Abid Naziri

mnaziri@ncsu.edu | github.com/abidnazirisami | abidnazirisami.github.io | linkedin.com/in/abidnazirisami

About Me

My research focuses on testing software, specifically learning-enabled systems. My current projects include testing deep learning libraries such as **PyTorch** and **TensorFlow** and simulation-based test generation for **Autonomous Driving Systems**. I have reported 94 bugs (74 confirmed) to **PyTorch** and **TensorFlow** to date [\[list\]](#).

Education

NC State University, PhD in Computer Science (4th year) May 2027 (Expected)

Advisor: Dr. Marcelo d'Amorim

Award: Graduate Merit Award - Summer 2025

Publications

BugsInDLLs: A Database of Reproducible Bugs in Deep Learning Libraries to Enable Systematic Evaluation of Testing Techniques [ISSTA 2025](#)

M M Abid Naziri, Aman Kumar Singh, Benjamin Wu, Feiran (Alex) Qin, Saikat Dutta, Marcelo d'Amorim

- Developed an open-source dataset **BugsInDLLs** with 112 reproducible bugs across three popular DL libraries: [JAX](#), [Tensorflow](#), and [PyTorch](#).
- Added support for benchmarking other fuzzing tools with an included integration of [FreeFuzz](#).

Venue: [The ACM SIGSOFT International Symposium on Software Testing and Analysis \(ISSTA 2025: Tool Demonstration Track\)](#)
[\[pdf\]](#)

Evaluating the Effectiveness of Neurosymbolic Constraint Learning for Testing Deep Learning Library APIs Submitted

M M Abid Naziri, Shinhae Kim, Feiran Qin, Saikat Dutta, Marcelo d'Amorim

- Proposed a tool called **Centaur** that uses constraints for DL APIs to generate valid test inputs.
- Uncovered 23 new bugs in [Tensorflow](#) and [PyTorch](#) (11 confirmed).

Testing Autonomous Driving Systems with Focused Misbehavior Forecasting Submitted

M M Abid Naziri, Stefano Carlo Lambertenghi, Andrea Stocco, Marcelo d'Amorim

- Developed **Foresee**, a tool to enhance simulation-based testing of self-driving software by identifying and analyzing near-miss scenarios, improving failure detection efficiency and effectiveness.
- **Foresee** exposes 128% more failures than state-of-the-art failure predictors.
- **Foresee** also enhances the capability of an existing technique [DriveFuzz](#) by upto 94%.

Evaluating the Effectiveness of Coverage-Guided Fuzzing for Testing Deep Learning Library APIs Submitted

Feiran Qin, *M M Abid Naziri*, Hengyu Ai, Saikat Dutta, Marcelo d'Amorim

- Conducted a study to assess the effectiveness of coverage guided fuzzing to test Deep Learning Library APIs.
- Reported 42 new bugs in [Tensorflow](#) and [PyTorch](#) (8 Fixed).

[\[preprint\]](#)

Evaluating the Effectiveness of Machine Learning to Improve Deep Learning Library Testing Submitted

Facundo Molina, *M M Abid Naziri*, Feiran (Alex) Qin, Alessandra Gorla, Marcelo d'Amorim

- Applied ML classifiers to infer the validity of an input before executing to increase efficiency of DL Library fuzzing tools.
- The classifiers achieved an accuracy of 91% in predicting input validity.
- They also improved an existing fuzzing tool called [ACETest](#) by increasing its validity ratio from 29% to 61%.

Experience

Senior Software Engineer, [Enosis Solutions](#) – Dhaka, Bangladesh

Mar 2019 – Jul 2022

- Developed scalable web applications using Vue JS, TypeScript, and ASP.NET Core, including a project management tool deployed to **270** employees and a Welder Management system
- Developed desktop applications DesignCalcs and ProWrite for pressure vessel design, adding features, fixing bugs, and creating unit/regression tests
- Managed CI/CD pipelines, configured testing servers, and ensured smooth system operations
- Operated in an **Agile** development process in a development team of 7 people
- Mentored new recruits, conducted code/design reviews
- Collaborated with clients and team members via **JIRA**

Technologies

Languages: Python, R, C#, Delphi, Java, C/C++, JavaScript, TypeScript, Bash

Technologies: PyTorch, TensorFlow, Carla Simulator, ASP.NET Core, React JS, Vue JS, TestComplete, DUnit